SAGE Research Methods Cases

Medicine & Health

Case Title

Challenges and Joys of using the Experience Sampling Method to Examine the Impact of the Arts on Wellbeing

Author Name(s)

Nicola J. Holt

Author Affiliation & Country of Affiliation

Department of Health and Social Sciences

University of the West of England, Bristol, UK

Lead Author Email Address

Email: Nicola.Holt@uwe.ac.uk

Discipline

Public Health [D26]

Sub-discipline

If you chose Medicine, Public Health, or Nursing as discipline, pick sub-discipline from the relevant list below. There are no sub-discipline options for Dentistry.

[For Medicine: choose sub-discipline]

Epidemiology & Biostatistics [SD-PH-2]

[For Nursing: choose sub-discipline]

Academic Level of intended readership

Postgraduate

Contributor Biographies

Dr Nicola Holt is a Senior Lecturer in Psychology in the Department of Health and Social Sciences at the University of the West of England, Bristol. She has written numerous articles on the topics of art, creativity and mental health and has an interest in research methodologies to investigate conscious experience. Her recent research has focused on the arts and health, and she is currently using repeated sampling methods to track participants' wellbeing across the duration of arts on prescription courses.

Published Articles

Holt, N. J. (2018). Using the experience-sampling method to examine the psychological mechanisms by which participatory art improves wellbeing. *Perspectives in public health*, *138*(1), 55-65.

Holt, N. J. (2019). The expression of schizotypy in the daily lives of artists. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(3), 359-371.

Abstract

This case study describes how the experience sampling method can be used to explore the impact of everyday behaviours on health and wellbeing, focusing on the correlates of participatory art in everyday life. The experience sampling method repeatedly samples moments of immediate experience as people go about their daily lives, tracking behaviours and their correlates, enabling temporal and complex patterns to be explored. The advantages of this compelling method are outlined, and practical advice given about how to run an experience sampling study successfully. For example, how to develop research questions at different levels of complexity, how to write an experience sampling questionnaire and how to engage and retain research participants.

Learning Outcomes

By the end of this case students should be able to:

- Develop research questions that draw on the specific advantages of the experience sampling method
- Understand the challenges in designing an experience sampling questionnaire and sampling schedule
- Consider approaches to engaging, and encouraging high response rates from, participants
- Show awareness of the recommendations for data screening and analysis for experience sampling research

---

Case Study

Project Overview and Context

As an early career researcher, I was interested in the role of the arts in improving wellbeing. While research supporting the benefits of the arts for health was accumulating, there was little focus on the mechanisms to explain this (Stuckey & Nobel, 2010). I had come across the experience sampling method through the work of Csikszentmihalyi (e.g. Csikszentmihalyi & Larson, 1987), and early work developing this method to explore correlates of stress and psychopathology in everyday life (e.g., de Vries, 1992). The method compelled me due to the richness of data that can be collected and the complex research questions that it enabled. I decided to use the experience sampling method to better understand experiences associated with art-making in everyday life and how these might relate to wellbeing.

Before I explain further why I chose the experience sampling method (ESM), it would be useful to describe it. ESM (sometimes known as ecological momentary assessment; EMA) is a structured, quantitative diary technique that enables the collection of longitudinal data. Participants complete 'experience sampling questionnaires' (ESQs) on multiple occasions, reporting on their 'immediate experience' as they go about their daily lives. Reports are usually triggered at either random or fixed intervals over a set period-of-time (signal-contingent sampling) (e.g., signalled by a 'beep' on one's mobile device;

Berkel, Ferreira & Kostakos, 2017). Although, reports can be triggered by a particular environmental, behavioural or psychological stimulus, such as smoking a cigarette or having a migraine (event-contingent sampling). Thus, ESM is described as a method that collects "multiple snapshots" of contextualized experience (Conner, Barrett, Tugade & Tennen, 2009).

Despite the ESM not previously having been used in the context of the arts and health, I chose this method for several reasons. Firstly, ESM has high ecological validity. Participants provide self-reports about the nature and quality of their experience as they go about their usual activities, making reports in a 'natural setting'. The reporting of events in this way recognises the importance of the contexts in which psychological and physiological processes unfold, enabling the mapping of thoughts, feelings and actions outside the laboratory and within the context of everyday life. Hence, in terms of researching art and health, rather than imposing a pre-selected and uniform art activity in an experimental trial, the impact of individuals' preferred and usual form of art-making, in their own art-making space, and at their own chosen time, could be examined, allowing naturally occurring triggers for, and consequences of, art-making to be explored.

Secondly, participants make reports on repeated occasions. This enables patterns over time to be examined. ESM can explore experiential (or temporal) regularities on three levels: 1) variation within an individual's experience (the 'experiential' level), for example, how do experiences of pain fluctuate in daily life, and what factors predict this?; 2) differences between groups of individuals based on average experiences (the 'person level'), for example, do certain coping styles (e.g. pain acceptance) predict less attention to pain in daily life?; and 3) the interaction between the person and experiential levels, for example, do people with certain coping styles experience less pain and more positive affect when carrying out everyday tasks? I was interested in the complexity that this repeated data enabled and excited by the opportunity to develop and test complex hypotheses, which hadn't been explored before in the context of the arts and health.

A final compelling feature of the ESM is that people report on their immediate experience – reports are made in 'real time'. This is thought to minimise problems associated with the fallibility of memory in the reconstruction of events. In terms of measuring wellbeing, this is especially important, since subjective wellbeing, is notoriously difficult to measure (accurately), and retrospective appraisals of wellbeing are prone to reporting errors and biases (Dolan & Metcalfe, 2012). For example, we tend to forget how we have been feeling over the course of a week, and we make systematic errors in our recall, for example, being biased towards remembering more positive affect (Ben-Zeev, Young & Madsen, 2009). ESM attempts to limit some of these problems by only asking people to focus on what was happening when 'you heard the beep'.

Section summary

- The ESM has three broad characteristics: participants record information in a 'natural setting'; in 'real time' (i.e. as close as possible to a signal or event); and do so on repeated occasions (often randomly sampling experience)
- The ESM enables interesting and complex research questions to be explored, since it tracks experiences and behaviours, and their correlates, over time
- I chose this method to help develop understanding of how art-making in everyday life can improve wellbeing

Research Design

Developing research questions

Once I had decided to use the ESM, designing the study was more challenging. Firstly, I needed to hone my research questions, based on the two levels of data described above (experiential and person levels) and their interaction. The experiential level consists of the repeated reports of individuals, and questions focused on this level can examine patterns over time (e.g. diurnal mood variation) or associations between different aspects of concurrent experience (e.g. context and mood). Drawing on three models of the relationship between creativity and mental health (Holt, 2019), my first question asked whether art-making in everyday life would predict improvements in mood and self-esteem, absorbed states of consciousness (the flow state), and cognitive changes (e.g., vivid internal imagery and playful thinking, akin to daydreams). The person level enables the grouping of experiential data according to individual characteristics (e.g. age, gender or scores on stable psychometric tests), so that comparisons can made, for example, of average levels of mood, or variance of mood (fluctuation), across a week. At this level, I asked whether people who made more art (on average) would score more highly on wellbeing measures. But, the most interesting research questions related to patterns across levels (cross-level interactions). Here, one can ask more nuanced and complex questions, that look at experiential patterns (rather than averages) according to grouping variables. I was interested in exploring whether experiences associated with art-making predicted wellbeing. For example, whether shifting into a positive mood after art-making predicted wellbeing, or whether wellbeing might be better explained by absorbed, concentrative states while art-making. This was interesting to me, because, although the ESM is observational, here we can begin to explore patterns over time that infer mechanisms to explain why art-making might improve wellbeing.

Choosing target variables

Subsequent design decisions, required at the start of any ESM project, included what sample to work with, which target variables to include, and what sampling strategy to use (when to signal participants). In terms of target variables, I decided to take a broad approach to the measurement of wellbeing (at the 'person level'), asking people to complete several self-report measures about how they generally think and feel, in order to include outcomes that had been used in previous art and health studies, including

happiness, self-esteem and self-regulation (Camic, 2008; King, 2001; Ullrich & Lutgendorf, 2002). However, a more difficult challenge was to design the experience sampling questionnaire, measuring 'in-the-moment' experience.

Writing the Experience Sampling Questionnaire (ESQ)

There was less previous research to guide the choice of 'in-the-moment measures'. Obvious state measures to include were mood and the flow state, which had been implicated in the arts (e.g., Csikszentmihalyi, 1996). However, I was interested in exploring conscious experience more broadly, to include cognition, such as internal imagery, which may play a role in self-regulation. As such, I needed to develop my own ESQ. I needed psychometrically validated 'in-the-moment measures' that were short and suitable for quick, repeated use. This proved challenging, for three reasons: existing ESQs were often not psychometrically validated; were too long, perhaps focusing on one aspect (e.g. mood) in depth; or the language was skewed (mainly towards psychopathology, e.g. focusing on cognitive intrusions associated with psychosis). After reviewing and rejecting many sub-scales, I managed to build an ESQ with 25-items, akin in length to ESQs used in previous research (e.g. Hektner, Schmidt & Csikszentmihalyi, 2007, pp. 293-297). I built this around a standard ESQ format, starting with open questions that set the context (e.g., 'where are you?'), before moving on to more specific experiences, with closed answers. This part of the ESQ was 'pieced together', using items from pre-existing ESQs that asked about the flow state and self-esteem, short mood checklists, and items from the Phenomenology of Consciousness Inventory (PCI; Pekala, 1991). The PCI was developed to comprehensively assess the phenomenological features of conscious experience, using neutral language, which was conceptually ideal. However, it had 53 items, which was far too many, and I chose subscales from this very carefully, focusing on those that measured daydream-like cognition. However, the most difficult and important aspect of the ESQ related to tracking art-making, since this had not been done before.

How to track art-making in everyday life – the sampling schedule

Art-making was the key variable in the study. I needed to consider how to best track this over time. This brought up wider issues regarding the design of the sampling schedule – when and how often to trigger participants' reports. The sampling schedule outlines how reports are triggered (e.g. by a particular event or a random stimulus), as well as many reports should be made each day and for how many days or weeks. Previous signal-contingent designs have had between four and ten signals a day – six being average – and have typically run for three days to three weeks (Conner & Lehman, 2012). Initially, I considered whether an 'event-contingent' sampling design would be appropriate, asking people to complete the ESQ whenever they 'felt like making art' or had 'just finished making art', for example. However, this was problematic since people may not wish to, or remember to, complete an ESQ at a point when they are intrinsically driven to do something else. Event-contingent designs can introduce selection biases, where, for example, certain art-making experiences may be more likely to be reported on, and may therefore be over-represented. Further, with this type of design reports of 'non-art-making' experiences would not be available for the purposes of comparison. For these

reasons, I chose a random signal-contingent sampling schedule. By so doing I hoped to obtain a representative sample of typical conscious experience, and within this, to have a sub-sample of experiences relating to art-making. However, how could I track art-making? I needed to do this with a question on the ESQ, that made sense within the broader sampling schedule.

I tracked art-making, by asking at each signal: "Since you were last beeped how much time have you spent making art?" (using a sliding response scale, with scores ranging from 0 to 100, indicating 'no time' to 'all of the time'). This strategy required relatively frequent signals each day (without becoming too interruptive and burdensome), making it likely that experiences of art-making would be 'captured'. I chose ten signals per day to allow this, and, in order to reduce overall participant burden, sampled experience for one week only. In order to be sure that this strategy would work, I asked participants to choose a week to collect data in which they would be making art.

What population to work with?

An important decision was who to work with, and how to define 'art'. I decided to work with people who self-defined themselves as artists, either as a hobby, semi-professional or professional. Artists were required to regularly engage in the arts, to have been doing so for at least one year and to be making art during the sampling week. This definition fit my purpose, being interested in how people use art in everyday life, perhaps without even realizing it, to improve their wellbeing. The one-year requirement was included so that familiarity with this art practice and way of working had been established.

Section summary

- Clear research questions, that relate to different levels of analysis (person and experience) and their interaction, are crucial
- A short, psychometrically valid, Experience Sampling Questionnaire is required to repeatedly measure target variables and their contextual predictors
- The sampling schedule needs to be planned carefully and be suitable for the type of experience or behavior under study – outlining how often and when to solicit reports on experience

Research Practicalities

Choosing the technology platform

Technology platforms serve the function of displaying the ESQ questions, storing the completed data, and signaling participants when to make reports (e.g. by an 'app' on a mobile phone). Any ESM program needs to facilitate the chosen sampling strategy, appropriate question types and enable methodological features, such as control over the timing and completion of reports, allowing compliance to protocols to be confirmed. At the time of designing my study, the software that best met my needs was the Experience

Sampling Program and it ran on handheld computers (personal data assistants; PDAs) (Barrett & Barrett, 2007). This software had been carefully designed for different ESM sampling schedules and question response types, including my required random-signal design. But crucially, it had additional methodological features. These included: 1) enabling the PDA to be in 'sleep mode' when participants were not completing an ESQ; 2) enabling a 'time window' in which signals must be responded to, so that ESQs could not be completed half an hour later (rather than at the randomly selected time); and 3) recording how long participants took to answer each question on the ESQ. In my study participants had a three-minute window to respond to a signal, each ESQ was 'time stamped', and response times for answering each question were recorded. ESP also had a useful feature that allowed participants to personalize aspects of the schedule after the PDA was programmed. They could choose the time frame in which the beeps fell each day and select a 'beep' and volume that worked best for them.

I also needed a platform that allowed me to build different types of question and to control for order effects. Varying question types can reduce boredom and automatic responses, but also, some of the psychometrically validated scales required check boxes or sliding scales. ESP was flexible and included various question response formats, including adjective lists, where participants could 'check' words that described their current mood. It also enabled sensitive question randomization. I programmed the contextual, scene-setting questions to always come first, followed by the question tracking art-making. But, all subsequent questions were presented in a random order, to reduce anticipation and boredom effects.

Since ESP worked on PDAs I investigated various models and chose the Palm Zire 72 since it had a large coloured screen, a good battery life, and adequate memory capacity (to save the data from 70 ESQs). I purchased a 'fleet' of PDAs. Budgets need to be considered carefully, to ensure that the fleet is large enough for the purposes of the study. I purchased six PDAs, along with memory cards, protective cases, chargers, and extra styluses in case any got lost. I calculated that I would run 3-4 trials each week over the course of a three-month-long period, and that six PDAs should facilitate this. All that was left to do was to program them with the ESQ and sampling schedule and pilot them.

Participant recruitment and retention

Recruiting participants was relatively simple. I created flyers with which to advertise the study, which were emailed to local groups of artists, posted on local (Bristol) online forums for artists, and pinned on the notice boards of local art galleries. The flyers had links to a website that described the study, with a contact form to express interest in taking part. I planned to recruit 40 people, and this method was sufficient for this purpose. I was more concerned with retention.

I needed to find people willing to be on call for one week, willing to complete surveys during every waking day of this week, at any time, and to meet with me in person, twice. For the research data to be usable, I needed them to be 'compliant', missing as few

signals as possible. It is recommended that, for data to be representative, at least 50% of ESQs be completed (Conner & Lehman, 2012), but I wanted to aim higher than that. I needed a strategy not only to recruit participants, but to engage them. After reading about other researchers' experiences (e.g. Conner et al., 2009; Scollon, Prieto & Diener, 2009), I decided upon a three-pronged approach: 1) to recruit people to take part as co-researchers; 2) to build a good rapport with them; and 3) to remunerate them financially.

Due to the time investment that was involved, I invited people to take part in the study as co-researchers – where they were the expert in their own experience and had responsibility for collecting data on it. I hoped that this would motivate participants and make them feel valued, emphasising the importance of their involvement. Building a good rapport with participants is important in experience sampling research (Mehl & Conner, 2012; Scollon et al., 2009). I met with each person who was interested in taking part, and spent 1-2 hours with them chatting about the aims of the study and what was involved in taking part. I spent time listening to their own interests in art and psychology, and we built a shared understanding that the research study was important in some way. We spent time discussing the procedure of the study, practiced using the technology platform, and made sure that the questions made sense to them. We also made sure that the volume of the signal was set at their preferred level, and set the timing of the signals for a twelve-hour period that would best fit with their working day (most participants chose between 9 am and 9 pm). Finally, we arranged a time at their convenience to meet after the data collection to discuss their experience.

I offered participants £70 each for taking part in the study, to reward them for their cooperation. In terms of payment, this could be equated to £1 per signal (since there were 70 signals for each person), reinforcing the idea that each beep was important, but participants were paid the same amount irrespective of any missed data. I hoped that this amount would encourage participation without being overly coercive, preventing people taking part due to the financial reward alone (Conner & Lehman, 2012).

Overall, this approach appeared to work. 89% of beeps were responded to (response rates ranging between 64% to 97%). This was above the average rate of 76%, and was therefore deemed adequate (Larson, 2019). However, I had a relatively short sampling period – of only one week. For longer sampling periods further reflections on how to keep participants motivated may be required, including regular contact to help them feel part of a research team.

Ethical considerations

As already discussed, there is an ethical obligation when designing the study to try not to overburden participants, with, for example, a taxing sampling schedule. Pilot testing can help determine this. Although ESM studies can seem quite intrusive, participants seem to habituate to the process (Conner & Lehman, 2012). Nevertheless, it is important that participants are aware that they are free not to answer a specific question, or respond to a signal at a particular time, and that they can stop participation at any time with no

questions asked. I made it clear in the initial meeting that it was fine not to answer a stimulus if it was inappropriate at that time. I did not expect them to not go swimming or drive for a long period of time because they were worried about missing a signal. I reassured them that I expected them to miss a few. A further consideration is that completion of the ESQs may cause reflection on personal experiences, and participants may need space to discuss this at the end of the project. A skilled debriefing session is important, allowing insights and experiences to be shared as appropriate. Researchers may need to provide links to community resources for further support in case this is required.

Screening the data

Once the data have been collected they need to be downloaded, organized, scored and screened. Data needed to be examined to ensure participant compliance. If the response rate for a participant is less than 50% then it is recommended that their data be removed, since it is not thought to be representative of their experience (Mehl & Conner, 2012). Individual ESQs also need to be checked. Questions that have been completed extremely quickly (< 1 second) were assumed to be completed inaccurately, for example, by inadvertently skipping a question (Mehl & Conner, 2012). Further, if participants had completed less than 50% of questions in an ESQ properly (e.g., had left most items on a sliding scale at their default point), responses to this ESQ were not included in the analysis.

Learning new statistical methods

Experience sampling data is nested. After screening the data, I had 2495 reports on immediate experience, nested within 41 participants. Data that is hierarchical in this way requires the use of multi-level modelling in order to account for the lack of independence of data at the stimulus level (Bolger & Laurenceau, 2013). Methods such as repeated measures analysis of variance do not allow for this structure, nor for the inevitable missing data that occurs at different time points for each participant (Hox & Roberts, 2011). I read text books, went on a multi-level modelling course at the University of Bristol, and practiced lots, with online guides and data sets. This was useful to do before data collection, to help me specify cross-level research questions with confidence, but also for clarifying my thinking about statistical power. Initially I had planned to have 30 participants, but realized that while $n = 70$ at the experiential level was adequate, $n = 30$ at the person level was rather small, and so I increased this to 40.

Section summary

- It is important to choose a technology platform that facilitates the sampling schedule, but also has features to ensure the reliability of ESM data
- A strategy to optimize participant engagement is key
- Ethical considerations include considering the burden of participation in the study for participants, and ensuring the sampling schedule is manageable

- Data considerations – consider in advance how you will screen and analyse data

Method in Action

Trial co-ordination

Collecting data was enjoyable, I met interesting participants and discussed some of my favourite topics with them. However, it was time consuming. I spent days, each week, for a period of 3-4 months, briefing and debriefing participants. At times the co-ordination of trials, based on the next availability of a PDA, became rather complex. Although the data collection largely went to plan, a larger fleet of PDAs would have been useful to enable more flexibility with trial timings. For example, I would have had to worry less about the potential impact of participants rearranging debriefing sessions on future trials, since I needed their PDA for the next trial.

Technical difficulties

On two trials the PDA 'froze' and I needed to help participants in the middle of their sampling week. On one occasion a quick chat over the phone resolved the issue and the trial continued. On another occasion the program needed to be restarted and the data was lost. The participant kindly asked to start again. Despite these limited 'disasters', the PDA option worked well. My main concern was the PDAs lost all data if their batteries went flat. I was especially worried about delays in returning the PDA at the end of the trials leading to data loss. However, luckily, this did not occur (possibly due to reminders to charge the PDAs).

Managing large data sets

What took longer than anticipated was cleaning and screening the data. This was because one of the questions stored data in an unusual format, and my word of caution here is to check such details in advance. For the mood scales I used a checklist format. The responses were recorded in base code ten, and I needed to convert this to base code two. For example, if there were 4 mood adjectives, '15' might be recorded (base code ten), which, in base code two, is '1111', indicating that all 4 mood adjectives were checked. In total there were 9980 responses that needed to be converted in this way. Screening responses for other issues, discussed above, was likewise time consuming, and this needs to be adequately anticipated in the research project's time scale.

Piloting the ESQ

One or two participants mentioned that they found a particular question from the PCI, that asked about altered states, strange, since they hadn't had this particular experience. This question was part of a 4-item sub-scale, and perhaps I could have measured the relevant construct with fewer items. Another participant commented that the ESQ didn't enquire about embodied experience, for example, feeling mindfully aware of one's body in movement, which was an important aspect of their experience of art-making. These discussions made me wonder whether some co-production of the ESQ, or at least piloting

of it with artists, would have been useful to help include the most relevant questions to reflect their experience.

Section summary

- Think about the ESQ from the participants' perspective – for example, are the questions relevant and easy to understand
- Be prepared to be on call to troubleshoot in case of any technical difficulties, and try to avoid these through clear communication and instructions
- Create a realistic time frame to collect and screen data – allow time for potential set backs

Consider how data is stored on the technology platform and how you will access and format it, in advance

Practical Lessons Learned

Dealing with complex data

ESM data is rich and nested. Having clear and testable hypotheses helps to navigate the complexity of data that is accrued. Ideally, these will focus on the experiential level and cross-level interactions to get the most out of the method. Initially, I found the development of the latter challenging, but how to go about this became clearer after reading several research papers on experience sampling and learning about statistical approaches, including multi-level modelling. It was when I fully understood how I was going to test my hypotheses statistically that the entire process clicked into place, like a completed jigsaw.

Psychometrics

Writing a good ESQ requires a lot of thought and reflection. The ESQ needs to balance the needs of the researcher and the participant. It needs to measure what one wants to learn about as a researcher, while not having too many questions, so that it is not time consuming and invasive. A participant is more likely to respond to an ESQ upon hearing a 'beep' while cooking dinner or painting a picture if it is quick, easy and enjoyable to complete. Hence, sacrifices may need to be made – only including the most essential questions and constructs. At the person-level, it is useful to think practically about how the careful addition of psychometric measures might broaden one's research scope, making use of the rich experiential data that will be collected to explore correlates of it in more than one research paper.

Building relationships with participants

Experience sampling studies are time intensive. Researchers need to be prepared for this investment of time and effort. Yet, it may be an even greater challenge to find research participants willing to likewise invest time and effort. I was lucky to have participants who engaged with the study – as co-researchers they collected the data, and did so

conscientiously. Thinking about how to recruit and engage participants is crucial to the success of an ESM study.

Changing platforms

The technology available for the design of ESM studies has changed in numerous ways since the method's inception. Initially, researchers relied upon paper and pencil questionnaire booklets, to be completed when a pager or wrist watch made a signal. Available technologies are in flux and have changed much in the last few years and the interested researcher will need to investigate current equipment and software available for developing a valid ESM study (Berkel et al., 2017). There are now several new technology platforms available, with different phone, network, cost and sampling and design implications. For example, more recent software, using mobile phones has the advantage of storing data online where it is less likely to be lost. Careful consideration needs to be made of which platform would work best for you, your participants and your project. Different platforms require different degrees of financial support, so budget is also a consideration here.

Section summary

- Evaluate the currently available technology platforms and their unique features – choose the most appropriate one for your sample, hypotheses and budget
- Develop your survey design skills to help build a good ESQ
- Learn how to manage and analyse complex data before collecting it, to help design the best study

Conclusion

The experience sampling method was extremely useful for capturing naturally occurring incidents of a potential health promoting behaviour and examining the subsequent impact of this on mental wellbeing. The analyses revealed some interesting outcomes, suggesting, for example, that people with lower levels of overall wellbeing had the greatest increase in mood following art-making, implying that certain groups may benefit from the arts as a wellbeing intervention more than others. It was also useful for inferring mechanisms, for example, supporting affective models. However, ESM is only useful if the research questions one is interested in requires the observation of real-life complexities. The experience sampling method is best used to understand the unfolding of behaviours and experiences over time. If I was designing this project again, I would develop research questions that deepened understanding of this, modelling complex processes of change over time, conducting growth curves using multi-level analysis. This approach has potential, too, for exploring the trajectory of experience across public health wellbeing interventions. For example, ESM could be useful to track wellbeing over time, even before, during and after an arts for health intervention, to explore the impact of this on behaviour and wellbeing. For example, exploring how long any immediate 'boosts' in mood last for. There is also the tempting idea of combining self-report with other data,

including physiological data, in future work, which has its unique challenges (Mehl & Conner, 2012). Overall, I found ESM to be an extremely useful method for exploring how events and experiences unfold in real life. However, it is time consuming, and it might serve one well to reflect on Larson's (2019, p. 551) warning: "Don't do it. Don't do it unless you're willing to fully commit yourself, ESM is not just a method you can tag onto another study. *It is a way of life*".

Section summary

- ESM is a flexible and powerful method for exploring changes in wellbeing and health behaviours over time
- It comes with a warning – it is time consuming.

Classroom Discussion Questions

What kinds of research question might ESM be used to explore?
What are the advantages of asking people to report on their immediate experience?
What are the advantages of repeatedly measuring experience?
What contextual factors could be measured, and how, to learn more about the correlates of experience?
What might facilitate 'compliance' when participating in an experience sampling study?

Multiple Choice Quiz Questions

How many signals per day are typical of ESM studies?

A 10 - 20

B 2 - 16

C 4 – 10 CORRECT

What are the main methodological strengths of ESM?

A The data is self-report, repeated and complex

B The data is real, repeated and random CORRECT

C The data is complex, repeated and real

Why is multi-level modelling best suited to ESM data?

A The data is nested (e.g. repeated observations for N people)

B There is no missing data across the observation points

C Both of the above CORRECT

Declaration of Conflicting Interests

The Author declares that there is no conflict of interest.

Further Reading

Mehl, M. R., & Conner, T. S. (2012). (Eds.) *Handbook of research methods for studying daily life*. New York: Guildford Press.

Hektner, J. M., Schmidt, J. A. & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. London: Sage Publications.

Bolger, N., & Laurenceau, J-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York: Guildford Press.

Larson R., & Csikszentmihalyi M. (2014). The experience sampling method. In M. Csikszentmihalyi (Ed.), *Flow and the foundations of positive psychology* (pp. 21-34). Dordrecht: Springer.

Web Resources

https://methods.sagepub.com/reference/encyc-of-research-design/n140.xml

https://academy.pubs.asha.org/2014/11/experience-sampling-method/

References

Van Berkel, N., Ferreira, D., & Kostakos, V. (2018). The experience sampling method on mobile devices. *ACM Computing Surveys (CSUR)*, *50*(6), 93.

Barrett, D. J., Barrett, L. F. (2007). ESP, the experience sampling program. http://www.experience-sampling.org (accessed 15 March 2012).

Ben-Zeev, D., Young, M. A., & Madsen, J. W. (2009). Retrospective recall of affect in clinically depressed individuals and controls. *Cognition and Emotion*, *23*(5), 1021-1040.

Camic, P. M. (2008). Playing in the mud: Health psychology, the arts and creative approaches to health care. *Journal of Health Psychology*, *13*(2), 287-298.

Conner, T. S., & Mehl, M. R. (2012). *Handbook of research methods for studying daily life.* New York: Guilford Press.

Conner, T. S., & Lehman, B. (2012). Getting started: Launching a study in daily life. In M. R. Mehl and T. S. Conner (Eds.), *Handbook of Research Methods for Studying Daily Life* (pp. 89 – 107). New York: Guilford Press.

Conner, T. S., Tennen, H., Fleeson, W., & Barrett, L. F. (2009). Experience sampling methods: A modern idiographic approach to personality research. *Social and Personality Psychology Compass*, *3*(3), 292-313.

Csikszentmihalyi, M. (1006). *Flow and the psychology of discovery and invention*. New York: Harper Collins.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *Journal of Nervous and Mental Disease*, *175*(9), 526–536.

Dolan, P., & Metcalfe, R. (2012). Measuring subjective wellbeing: Recommendations on measures for use by national governments. *Journal of Social Policy*, *41*(2), 409-427.

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Thousand Oaks, CA: Sage.

Hox, J., & Roberts, J. K. (Eds.). (2011). *Handbook of advanced multilevel analysis*. New York: Routledge.

King, L. A. (2001). The health benefits of writing about life goals. *Personality and Social Psychology Bulletin, 27*, 798–807

Larson, R. W. (2019). Experiencing sampling research from its beginnings into the future. *Journal of Research on Adolescence*, *29*(3), 551-559.

Pekala, R. J. (1991). *Quantifying consciousness.* Boston, MA: Springer.

Scollon, C. N., Prieto, C. K., & Diener, E. (2009). Experience sampling: promises and pitfalls, strength and weaknesses. In E. Diener (Ed.), *Assessing well-being* (pp. 157-180). Springer, Dordrecht.

Stuckey, H. L., & Nobel, J. (2010). The connection between art, healing, and public health: A review of current literature. *American Journal of Public Health*, *100*(2), 254-263.

Ullrich, P. M., & Lutgendorf, S. K. (2002). Journaling about stressful events: Effects of cognitive processing and emotional expression. *Annals of Behavioral Medicine, 24*(3), 244-250.

de Vries, M. W. (Ed.) (1992). *The experience of psychopathology: Investigating mental disorders in their natural settings*. Cambridge: Cambridge University Press.